

The 27th International Conference on Database Systems for Advanced Applications (DASFAA-2022), April 11-14, 2022, Hyderabad, India. <https://www.dasfaa2022.org>

(Note: DASFAA-2022 conference is held in online mode.
All timings are as per Indian Standard Time (IST) (UTC + 05:30)

I. Workshop Program Schedule (April 11, 2022, Monday)

April 11, 2022 (Monday) All timings are as per Indian Standard Time (IST) (UTC + 05:30) For more details, please visit respective workshop website available at the following page. https://www.dasfaa2022.org/Programme/workshops					
Time Duration	Hall 1	Hall 2	Hall 3	Hall 4	Hall 5
09.00-13.00	International Workshop on Blockchain Technologies (IWBT2022) Organizers: Sanjay Chaudhary (Ahmedabad University, India), Krishnasuri Narayanam (IBM Research, India)	Managing Air Quality Through Data Science (MAQTDS 2022) Organizers: Girish Agrawal (O.P. Jindal Global University, India), Jai Ganesh (Mphasis Ltd., India)	1st Workshop on Pattern mining and Machine learning in Big complex Databases (PMBD 2022) Organizers: Philippe Fournier-Viger (Shenzhen University, China), Mourad Nouioua (Harbin Institute of Technology Shenzhen, China), Hamido Fujita (Iwate Prefectural University, Japan), Lin Zhang (Tencent.com), Vincent S. Tseng (National Chiao Tung University, Taiwan)	The 8th International Workshop on Big Data Management and Service (BDMS 2022) Organizers: Xiaoling Wang (East China Normal University, China), Kai Zheng (University of Electronic Science and Technology of China), An Liu (Soochow University, China)	The 7th International Workshop on Big Data Quality Management (BDQM 2022) Organizers: Xiaou Ding (Harbin Institute of Technology, China), Xueli Liu (Tianjin University, China)
13.00-14.00	Lunch Break				
14.00-18.00	IWBT2022 (Afternoon session)	MAQTDS 2022 (Afternoon session)	The 6th International Workshop on Graph Data Management and Analysis (GDMA 2022) Organizers: Lei Zou (Peking University, China)		

II. Main Conference Program Schedule

(April 12-14, 2022, Tuesday to Thursday)

April 12, 2022 (Tuesday)					
All timings are as per Indian Standard Time (IST) (UTC + 05:30)					
Time Duration	Hall 1	Hall 2	Hall 3	Hall 4	Hall 5
08.30-09.00	Inauguration Lei Chen (Hong Kong University of Science and Technology, Hong Kong), P. J. Narayanan (IIIT Hyderabad, India), S. Sudarshan (IIT Bombay, India), Masaru Kitsuregawa (University of Tokyo, Japan), P. Krishna Reddy (IIIT Hyderabad, India), Mukesh Mohania (IIIT Delhi, India), Anirban Mondal (Ashoka University, India), Arnab Bhattacharya (IIT Kanpur, India), Lee Mong Li Janice (National University of Singapore, Singapore), Divyakant Agrawal (University of California, Santa Barbara, USA), Prasad M. Deshpande (Google), Rajeev Gupta (Microsoft, India)				
09.00-10.00	Keynote by Gautham Das (University of Texas at Arlington) (Hall 1) Title: Fairness in Database Querying Chair: Mukesh Mohania (IIIT, Delhi)				
10.00-10.15	Tea break				
10.15-12.30	Research Session #1 (Queries) Chair: P Radha Krishna (NIT Warangal, India)	Research Session #2 (Text and Image-I) Chair: Anil Kumar Vuppala (IIIT Hyderabad, India)	Research Session #3 (Applications of ML-I) Chair: <>	Industry Session #1 (Advancing Recommendation Systems) Chair: Prasad Deshpande (Google)	Tutorial Session #2 (Title: Reachability on Large-scale Graphs: Models, Techniques, and Trends) Chair: Lini Thomas (IIIT Hyderabad)
12.30-13.30	Lunch break				
13.30-14.45	Demo and PhD Consortium Chair: Anirban Mondal (Ashoka University, India)				
14.45-16.30	Research Session #4 (Graphs-I) Chair: Vasudha Bhatnagar (University of Delhi)	Research Session #5 (Text and Image-II) Chair: <>	Research Session #6 (Recommendation-I) Chair: Manish Singh (IIT Hyderabad)		Tutorial Session # 1 (Title: Make Wise Decisions for Your DBMSs: Workload Forecasting and Performance Prediction Before Execution) Chair: Vikram Pudi (IIIT Hyderabad)
16.30-17.00	Tea break				
17.00-18.30	Panel Session Moderator: Kurt Stockinger (Zurich University of Applied Sciences, Switzerland) Panelists: Georgia Koutrika (Athena Research, Greece), Jaydeep Sen (IBM Research, India), Immanuel Trummer (Cornell University, USA) Lei Zou (Peking University, China)				

April 13, 2022 (Wednesday)

All timings are as per Indian Standard Time (IST) (UTC + 05:30)

Time Duration	Hall 1	Hall 2	Hall 3	Hall 4	Hall 5
09.00-10.00	Keynote by Tirthankar Lahri (Oracle Corporation) (Hall 1) Title: Oracle Database In-Memory: The Enterprise, at Warp Speed! Chair: Masaru Kitsuregawa (University of Tokyo, Japan)				
10.00-10.15	Tea break				
10.15-12.30	Research Session #7 (Spatio-temporal Data) Chair: Uday Kiran, (The University of Aizu, Japan)	Research Session #8 (Recommendation-II) Chair: Vikram Goel (IIIT Delhi)	Research Session #9 (Applications of ML-II) Chair: <>	Industry Session #2 (Data Management and Search) Chair: Rajasekar Krishnamurthy (Adobe)	Tutorial Session #3 (Title: A tutorial on biomedical image segmentation using deep learning) Chair: Sudipta Banerjee (IIIT Hyderabad, India)
12.30-13.30	Lunch break				
13.30-14.30	Keynote by: Ioana Manolescu (Inria and Institut Polytechnique de Paris) (Hall 1) Title: Teasing journalistic findings out of heterogeneous sources: a data/AI journey Chair: Srinath Srinivasa (IIIT Bangalore, India)				
14.30-16.15	Research Session #10 (Algorithms-I) Chair: Anirban Mondal (Ashoka University)	Research Session #11 (Systems) Chair: Satish Narayana Srirama (University of Hyderabad)	Research Session #12 (Applications of ML-III) Chair: <>		Tutorial Session #4 (Title: AI Meets NoSQL Database: Methods, Opportunities and Challenges) Chair: Kamal Karlapalem (IIIT Hyderabad)
16.15-16.30	Tea break				
16.30-18.00	Research Session #13 (Security) Chair: Ashok Kumar Das (IIIT Hyderabad)	Research Session #14 (Text and Image-III) Chair: Anil Kumar Vuppala (IIIT Hyderabad, India)	Research Session #15 (Graphs-II) Chair: Tanmoy Chakraborty (IIIT Delhi)		


April 14, 2022 (Thursday)

All timings are as per Indian Standard Time (IST) (UTC + 05:30)

Time Duration	Hall 1	Hall 2	Hall 3	Hall 4	Hall 5
09.00-10.00	Keynote by Sunita Sarawagi (IIT Bombay) (Hall 1) Title: Modern AI for Age-old problems of Database Systems Chair: Sang-Won Lee (SKKU, Korea)				
10.00-10.15	Tea break				
10.15-12.30	Research Session #16 (Algorithms-II) Chair: Uday Kiran, (The University of Aizu, Japan)	Research Session #17 (Recommendation-III) Chair: <	Research Session #18 (Applications of ML-IV) Chair: <	Industry Session #3 (Industrial Machine Learning Applications) Chair: Daxin Jiang, (Microsoft)	Tutorial Session #5 (Title: Time Series Anomaly Detection Toolkit for AI Applications) Chair: Praveen Paruchuri (IIIT Hyderabad)
12.30-13.30	Lunch break				
13.30-14.30	Keynote by Guoliang Li (Tsinghua University, Beijing, China) (Hall 1) Title: openGauss: An Autonomous Database System Chair: Arnab Bhattacharya (IIT Kanpur, India)				
14.30-16.15	Research Session #19 (Knowledge Bases) Chair: V. Raghava Mutharaju (IIIT Delhi)	Research Session #20 (Text and Image-IV) Chair: <	Research Session #21 (Recommendation-IV) Chair: V.Ravi (IDRBT, Hyderabad)		
16.15-16.30	Concluding Remarks (Hall 1) Chair: P.Krishna Reddy (IIIT Hyderabad)				

Program Details

1. Keynote talks (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Date and Time (Hall 1)	Details
Keynote 1 (April 12, 2022, Tuesday, 09.00- 10.00, Hall 1)	 <p>Gautam Das, Director of Center for Artificial Intelligence and Big Data (CARIDA) and Database Exploration Laboratory (DBXLAB), University of Texas at Arlington.</p> <p>Title: Fairness in Database Querying</p> <p>Abstract:</p> <p>We are being constantly judged by automated decision systems that have been criticized for being sometimes discriminatory and unfair. In this talk, we focus on fairness issues that arise when users perform ad-hoc exploration of databases using commonly available querying mechanisms such as selection/range queries, ranking queries, top-k queries, etc. For example, a user may use such queries to retrieve suitable employment opportunities in a jobs database, dating partners in a matching website, or apartments to rent in a real estate database. We will discuss how such querying mechanisms can give sometimes give results that are discriminatory, and discuss approaches to detect, mitigate and prevent such scenarios from occurring. Our work represents some of the initial steps towards the broader goal of integrating fairness conditions into database query processing and data management.</p> <p>Biography:</p> <p>Dr. Das is the Associate Dean for Research, College of Engineering, a Distinguished University Chair Professor of Computer Science and Engineering, Director of the Center for Artificial Intelligence and Big Data (CARIDA), and Director of the Database Exploration Laboratory (DBXLAB) at UT-Arlington. Prior to joining UTA in 2004, he has held positions at Microsoft Research, Compaq Corporation and the University of Memphis. He graduated with a B.Tech in computer science from IIT Kanpur, India in 1983, and with a Ph.D in computer science from the University of Wisconsin, Madison in 1990. He is a Fellow of the IEEE and a member of the ACM.</p> <p>Dr. Das has published over 200 papers, many of which have appeared in premier data mining, database and algorithms conferences and journals. His work has received several awards, including the Communications of the ACM Research Highlights in 2021, ACM SIGMOD Research Highlights in 2019, IEEE ICDE 10-Year Influential Paper Award in 2012, ACM SIGKDD Doctoral Dissertation Award (honorable mention) in 2014 for his former student, and numerous other awards. He has presented keynotes and invited lectures, tutorials and courses at various universities, research labs, and conferences. He has been on the Editorial Board of the journals ACM Transactions on Database Systems and IEEE Transactions on Knowledge and Data Engineering. He has served in the organization roles of several major conferences, including as General Chair of ACM SIGMOD/PODS 2018.</p>

Keynote 2
(April 13, 2022,
Wednesday,
09.00-10.00, Hall
1)



Tirthankar Lahiri, Senior Vice President, Data and In-Memory Technologies,
Oracle Corporation

Title: Oracle Database In-Memory: The Enterprise, at Warp Speed!

Abstract:

In-memory computing is more than simply about speed. It enables a fundamental transformation in business processes. Just as air travel enabled more than just the ability to travel faster: It enabled a completely new global economy, reshaped politics, and transformed society. Oracle's Database In-Memory feature similarly enables not just faster analytics, but a fundamental rethinking and drastic simplification of the traditional analytic platform. Combined with Oracle's many converged database capabilities that bring together many data models and many workloads, and with Oracle's Autonomous Database platform that makes self-driving machine-learning powered databases a reality, Oracle Database In-Memory allows for the development of a new category of enterprise architectures, with significant reduction in cost and complexity, while providing unmatched performance for both transactional and analytic workloads.

Biography:

Tirthankar Lahiri is Senior Vice President of the Data and In-Memory Technologies area within Oracle Database. This includes the Oracle Database Engine (Transactions, Data formats, Indexes, Advanced Compression, Database In-Memory, the Database Filesystem, etc.), the Oracle TimesTen In-Memory Database, and Oracle NoSQLDB. Tirthankar has 26 years of experience in the database industry and has worked on a number of areas such as Performance, Scalability, Manageability and In-Memory architectures. He has 45 issued and several pending patents and a number of academic publications. He has a B.Tech in Computer Science from the Indian Institute of Technology (Kharagpur) and an MS in Electrical Engineering from Stanford University. He was in the PhD program at Stanford and his research included NUMA Operating Systems (the Hive project) and Semistructured Data (the Ozone project) before his PhD was superseded by his industrial career.

Keynote 3
(April 13, 2022,
Wednesday,
13.30-14.30, Hall
1)



Ioana Manolescu, Inria and Institut Polytechnique de Paris.

Title: Teasing journalistic findings out of heterogeneous sources: a data/AI journey

Abstract:

Freedom of the press is under threat worldwide, and the quality of information that people have access to is dangerously degraded, under the joint threat of non-democratic governments and fake information propagation. The press as an industry needs powerful data management tools to help them interpret the complex reality surrounding us.

Since 2018, I have been cooperating with journalists from Le Monde, France's leading newspaper, in devising tools for analyzing large and heterogeneous data sources that they are interested in. This research has been embodied in ConnectionLens, a graph ETL tool capable of ingesting heterogeneous data sources into a graph, enriched (with the help of ML methods) with entities extracted from data of any type. On such integrated graphs, we devised novel algorithms for keyword search, and combine them in more recent research with structured querying. The talk describes the architecture and main algorithmic challenges in building and exploiting ConnectionLens graphs, illustrated in particular on an application where we study conflicts of interest in the biomedical domain. This is joint work with A. Anadiotis, O. Balalau, H. Galhardas and many others. ConnectionLens Web site (papers+code): <https://team.inria.fr/cedar/connectionlens/>

This research has been funded by Agence Nationale de la Recherche AI Chair SourcesSay (<https://sourcessay.inria.fr>)

Biography:

Ioana Manolescu is a senior researcher at Inria Saclay and a part-time professor at Ecole Polytechnique, France. She is the lead of the CEDAR INRIA team focusing on rich data analytics at cloud scale. She is also the scientific director of LabIA, a program ran by the French government whereas AI problems raised by branches of the local and national French public administration are tackled by French research teams. She is a member of the PVLDB Endowment Board of Trustees, and has been Associate Editor for PVLDB, president of the ACM SIGMOD PhD Award Committee, chair of the IEEE ICDE conference, and a program chair of EDBT, SSDBM, ICWE among others. She has co-authored more than 150 articles in international journals and conferences and co-authored books on "Web Data Management" and on "Cloud-based RDF Data Management". Her main research interests algebraic and storage optimizations for semistructured data, in particular Semantic Web graphs, novel data models and languages for complex data management, data models and algorithms for fact-checking and data journalism, a topic where she is collaborating with journalists from Le Monde. She is also a recipient of the ANR AI Chair titled "SourcesSay: Intelligent Analysis and Interconnexion of Heterogeneous Data in Digital Arenas" (2020-2024).

Keynote 4
(April 14, 2022,
Thursday, 09.00 -
10.00, Hall 1)



Sunita Sarawagi , Institute Chair Professor, Computer Science and Engineering, IIT Bombay.

Title: Modern AI for Age-old problems of Database Systems

Abstract:

Modern deep learning methods are pushing the frontiers of many challenging problems in database systems. We will discuss state-of-the-art machine learning models that are providing record breaking accuracies on age-old tasks such as entity resolution, missing value imputation and natural language querying. We are also witnessing brand new capabilities that were not possible a few years back. We can perform entity resolution across heterogeneous, multilingual datasets via actively learned nearest neighbor indices, thereby eliminating the need for hand-designing blocking predicates. On multi-dimensional analytical datasets, we can now obtain joint distributions over thousands of interacting time series. Advances in pre-trained language models have significantly increased the capability of handling natural variations in parsing text input to SQL. In this talk we will go over the latest ML research that is enabling these capabilities, and present directions for future research.

Biography:

Sunita Sarawagi researches in the fields of databases and machine learning. She is institute chair professor at IIT Bombay. She got her PhD in databases from the University of California at Berkeley and a bachelors degree from IIT Kharagpur. She has also worked at Google Research (2014-2016), CMU (2004), and IBM Almaden Research Center (1996-1999). She was awarded the Infosys Prize in 2019 for Engineering and Computer Science, and the distinguished Alumnus award from IIT Kharagpur. She has several publications including best paper awards at ACM SIGMOD, VLDB, ICDM, NIPS, and ICML conferences. She has served on the board of directors of the ACM SIGKDD and VLDB foundation. She was program chair for the ACM SIGKDD 2008 conference, research track co-chair for the VLDB 2011 conference and has served as program committee member for SIGMOD, VLDB, SIGKDD, ICDE, and ICML conferences, and on the editorial boards of the ACM TODS and ACM TKDD journals.

Keynote 5
(April 14, 2022,
Thursday, 13.30-
14.30, Hall 1)



Guoliang Li, Tsinghua University, Beijing, China.

Title: openGauss: An Autonomous Database System

Abstract:

In this talk, I will present how to build an autonomous database system. I discuss how to integrate effective learning-based models into database systems to build learned optimizers (including learned query rewrite, learned cost/cardinality estimation, learned join order selection and physical operator selection) and learned database advisors (including self-monitoring, self-diagnosis, self-configuration, and self- optimization). I also propose an effective validation model to validate the effectiveness of learned models. I discuss effective training data management and model management platforms to easily deploy learned models. Finally, I will introduce our autonomous database system openGauss.

Biography:

Guoliang Li is a full professor and the deputy head of Department of Computer Science, Tsinghua University, Beijing, China. His research interests include large-scale data integration and cleaning, human-in-the-loop data management, machine learning for database, and database for machine learning. He is a general co-chair of SIGMOD 2021, demo co-chair of VLDB 2021, industry co-chair of ICDE 2022, and PC co-chair of DASFAA 2019. He is also an associate editor of VLDB journal and IEEE TKDE. He is a steering committee

member of IEEE TCDE and DASFAA. He received best paper awards (candidates) of VLDB 2020, ICDE 2018, KDD 2018, CIKM 2017 and DASFAA 2014. He received Early Research Contribution Award of VLDB and Early Career Award of IEEE TCDE.

2. Panel Session (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Panel Session
(April 12, 2022,
Tuesday,
17.00-18.30, Hall
1)

Title: Futuristic Data Interfaces

Overview:

Database systems have been around for more than four decades and have been widely used in academia and industry across the globe. In this panel we discuss the following two questions from various perspectives with four different internationally renowned database experts.

- What are the challenges of current database interfaces?
- What solutions do you propose to solve these challenges?

Moderator:



Kurt Stockinger (Zurich University of Applied Sciences, Switzerland)

Panelists

- Georgia Koutrika (Athena Research, Greece)



- Jaydeep Sen (IBM Research, India)



- Immanuel Trummer (Cornell University, USA)



- Lei Zou (Peking University, China)



Moderator: Kurt Stockinger (Zurich University of Applied Sciences, Switzerland)

Prof. Dr. Kurt Stockinger is Professor of Computer Science, Director of Studies in Data Science at Zurich University of Applied Sciences (ZHAW) and Co-Head of the ZHAW Datalab. His research focuses on Data Science with emphasis on Big Data, Natural Language Query Processing, Query Optimization and Quantum Computing. Essentially, his research interests are at the intersection of databases, natural language processing and machine learning. He is also on the Advisory Board of Callista Group AG and the International AIQT Foundation. Previously Kurt Stockinger worked at Credit Suisse in Zurich, Switzerland, at Lawrence Berkeley National Laboratory in Berkeley, California, at California Institute of Technology, California as well as at CERN in Geneva, Switzerland. He holds a Ph.D. in computer science from CERN / University of Vienna.

(i) Georgia Koutrika (Athena Research, Greece)

Title: Intelligent Data Assistants

Abstract: Data is considered the 21st century's most valuable commodity. Analysts exploring data sets for insight, scientists looking for patterns, and consumers looking for information are just a few examples of user groups that need to access and dig into data. Despite technological advances in the data exploration and data management domains, existing systems are falling behind in bridging the chasm between data and users, making data accessible and useful only to the few. A futuristic data interface would enable interaction with data using natural language, would understand the data as well as the user intent, would guide the user, and make suggestions, and altogether help the user leverage data for all sorts of purposes (from finding answers to questions to revealing patterns and finding solutions to problems) in a more natural way. These systems, which we call intelligent data assistants, require the synergy of several technologies and innovation in all these fronts, including natural language interfaces, data exploration, conversational AI, and data management

Bio: Georgia Koutrika is a Research Director at Athena Research Center in Greece. She has more than 15 years of experience in multiple roles at HP Labs, IBM Almaden, and Stanford. Her work emerges at the intersection of data management, natural language processing and deep learning and focuses on intelligent and interactive data exploration, conversational data systems, and user-driven data management. Her work has been incorporated in commercial products, described in 14 granted patents and 26 patent applications in the US and worldwide, and published in more than 100 papers in top-tier conferences and journals. Georgia is an ACM Senior Member and IEEE Senior Member. She is a member of the VLDB Endowment Board of Trustees, member of the PVLDB Advisory Board, member of the ACM-RAISE Working Group, co-Editor-in-chief for VLDB Journal, PC co-chair for VLDB 2023, co-EIC of Proceedings of VLDB (PVLDB). She has been associate editor in top-tier conferences (such as ACM SIGMOD, VLDB) and journals (VLDB Journal, IEEE TKDE), and she has been in the organizing committee of several conferences including SIGMOD, ICDE, EDBT, among others. She has received a PhD and a diploma in Computer Science from the Department of Informatics and Telecommunications, University of Athens, Greece.

(ii) Jaydeep Sen (IBM Research, India)

Title: Evolution of NLIDB systems and their application in Industry setups

Abstract: In this modern era of technology, multitude of business applications are rapidly moving towards data driven insights for intelligent decision making, analytics and more. As we continue to see heaps of digital exhaust being generated, access to data is still limited to technical users who can query the datastores with specific query languages. Natural Language Interface to

Databases (NLIDB) systems have gained a lot of focus recently owing to its fascinating aim of democratizing data access to non-technical business users. The research space for NL interfaces for data has evolved a lot since its inception, starting from simple keyword based queries, all the way to machine learning based systems, also dubbed as text-to-sql challenge. While the appeal of NLIDB system is common across different persona and use-cases, deploying a NLIDB system for an industry application has its own set of challenges which are often closely coupled with the exact domain and use-case. With no "one size fits all" solution in place, it is the right time for the community to review how the different methodologies adopted for NLIDB systems correlate with their applicability across different use-cases seen in academia and industry.

Bio: Jaydeep Sen is a Research Staff Member in IBM Research AI, India Lab . His research interests include applications for natural language understanding, semantic reasoning, designing intelligent algorithms for "learning from small data" applications. His work at IBM has powered some of IBM's most prominent QA and NL application portfolio. He has publications at conferences like VLDB, SIGMOD, IJCAI, IEEE SCC, EMNLP, COLING etc. and has served as Program Committee members for AAAI, SIGMOD, ICDE etc. He has more than 20 patents (granted/filed) in USPTO as of Dec-2021.

(iii) Immanuel Trummer (Cornell University, USA)

Title: Voice Interfaces for Data Access

Abstract: The communication between user and user is shifting more and more towards voice interfaces. This trend is evidenced by devices and services such as Google Home, Amazon Alexa, or Apple's Siri. For many users, speech is the most natural form of interaction. It enables computer use from a distance, even in scenarios where hands or the visual attention are bound (e.g., while driving). All those advantages motivate the question of how to leverage voice interfaces for convenient data access.

Accessing data via voice query interfaces is challenging. First, noisy speech recognition adds uncertainty on top of the inherent difficulties of natural language understanding. Second, transferring query results to users via voice output is difficult. Verbose speech output risks overwhelming the listener. Hence, output needs to summarize and to focus on the most important trends in the data. Recent research tackles some of those challenges. Still, many research questions remain open and must be answered to make the vision of natural data access via voice query interfaces a reality.

Bio: Immanuel Trummer is assistant professor at Cornell University, working towards making data analysis more efficient and more user-friendly. His papers were selected for "Best of VLDB", "Best of SIGMOD", for the ACM SIGMOD Research Highlight Award, and for publication in CACM as CACM Research Highlight. His current research is funded by the NSF and by multiple Google Faculty Research Awards.

(iv) Lei Zou (Peking University, China)

Title: Natural Language Question Answering over Knowledge Graph

Abstract: As more and more structured data become available on the web, the question of how end users can access this body of knowledge becomes of crucial importance. As a de facto standard of a knowledge base, RDF (Resource Description Framework) repository is a collection of triples, denoted as <subject, predicate, object>. Although SPARQL is a standard way to access RDF

	<p>data, it remains tedious and difficult for end users because of the complexity of the SPARQL syntax and the RDF schema. An ideal system should allow end users to profit from the expressive power of Semantic Web standards (such as RDF and SPARQLs) while at the same time hiding their complexity behind an intuitive and easy-to-use interface.</p> <p>Generally, there are two categories of existing methods on natural language question answering (Q/A) over RDF database---one is IR (Information Retrieval)-based and the other one is called semantic parsing method. In this panel, I will talk about our solution gAnswer, which is based on graph matching-based technique, to design an effective natural language interface to access KG database.</p> <p>Bio: Lei Zou is a professor at Peking University, China, and his recent research interests include graph databases, knowledge graph, particularly in graph-based RDF data management, natural language question answering over knowledge graph, hardware assisted graph database systems. Lei Zou's research is supported by multiple NSFC projects. Prof. Zou also obtained Newton Advanced Fellowships of UK Royal Society. Lei Zou has publications at conferences like VLDB, SIGMOD, ICDE etc, and has served as Program Committee members for SIGMOD, VLDB and ICDE. He served PC Area Chair of ICDE 2021 and PC Chair of WISE 2022. Now, he is an Associate Editor of IEEE Transactions on Knowledge and Data Engineering (TKDE).</p>
--	---

3. Tutorial talks (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Date and Time	Tutorial talks
Tutorial # 1 (April 12, 2022, Tuesday, 14.45-16.30, Hall 5)	<p>Title: Make Wise Decisions for Your DBMSs: Workload Forecasting and Performance Prediction Before Execution</p> <p>Speakers:</p> <ul style="list-style-type: none"> • Zhengtong Yan (University of Helsinki) • Jiaheng Lu (University of Helsinki) • Qingsong Guo (University of Helsinki) • Gongsheng Yuan (University of Helsinki) • Calvin Sun (Huawei Toronto) • Steven Yang (Huawei Toronto). <p>Brief outline of the tutorial:</p> <p>In this tutorial, we will focus on 1) how to forecast the future workloads (e.g., workload shift detection, arrival rate prediction, and next query prediction), and 2) how to analyze the behaviors of the workloads (e.g., execution time prediction and resource usage estimation). We will provide a comprehensive overview and detailed introduction of the two topics, from state-of-the-art methods, real-world applications, to open problems and future directions. Specifically, we will not only discuss traditional methods, such as time-series analysis, Markov modeling, analytical modeling, and experiment-driven methods, but also cover the state-of-the-art AI techniques, including machine learning, deep learning, reinforcement learning, and graph embedding.</p> <p>Biographies of Speakers:</p> <p>Zhengtong Yan is a doctoral student at the University of Helsinki. His research topics include autonomous multi-model databases and cross-model query optimization.</p> <p>Jiaheng Lu is a professor at the University of Helsinki. His main research interests lie in database systems specifically in the challenge of efficient data processing from real-life, massive data repositories and the Web. He has written four books on Hadoop and NoSQL databases, and more than 100 papers published in SIGMOD,</p>

	<p>VLDB, TODS, and TKDE, etc. He has given several tutorials on multi-model data management and autonomous databases in VLDB, CIKM, and EDBT conferences. He frequently serves as a PC member for conferences including SIGMOD, VLDB, ICDE, EDBT, CIKM, etc.</p> <p>Qingsong Guo is a postdoctoral researcher at the University of Helsinki His research interests include multi-model databases and automatic management of big data with deep learning.</p> <p>Gongsheng Yuan is a doctoral student at the University of Helsinki. His research topics lie in databases with quantum theory or reinforcement learning.</p> <p>Calvin Sun is the Chief Database Architect at Huawei Cloud. He has 20+ years of experience in developing several database systems, ranging from embedded databases, large-scale distributed databases, to cloud-native databases.</p> <p>Steven Yuan is the Director of Huawei Toronto Distributed Scheduling and Data Engine Lab. He leads a research team in the big data and cloud domain, focusing on distributed scheduling and distributed database, from IaaS to PaaS.</p>
<p>Tutorial # 2, Tuesday, (April 12, 2022, 10.15-12.30, Hall 5)</p>	<p>Title: Reachability on Large-scale Graphs: Models, Techniques, and Trends</p> <p>Speakers:</p> <ul style="list-style-type: none"> • Xiaoshuang Chen (Guangzhou University) • Long Yuan (Nanjing University of Science and Technology) • Wenjie Zhang (University of New South Wales) • Ying Zhang (University of Technology Sydney) <p>Brief outline of the tutorial:</p> <p>In this tutorial, we will first show the importance and challenges of studying reachability queries. Then, we will introduce the existing reachability models defined over different graphs. The computing algorithms regarding different models and environmental settings will also be presented. Finally, we will discuss future research directions in this important research area.</p> <p>Biographies of Speakers:</p> <p>Xiaoshuang Chen is an Associate Professor in the Cyberspace Institute of Advanced Technology, Guangzhou University. Before that, she was a Postdoctoral Fellow in the School of Computer Science and Engineering, University of New South Wales. Her research interest lies in large-scale graph data analysis. She has published several papers in ICDE, VLDB and VLDBJ since 2017.</p> <p>Long Yuan is a Professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. His research focuses on graph data management and analysis. He has published papers in top venues such as VLDB, WWW, ICDE, VLDBJ, and TKDE.</p> <p>Wenjie Zhang is an Australian ARC Future Fellow (2021-2025) and Professor in the School of Computer Science and Engineering, University of New South Wales. Her research interests include spatial-temporal data analysis and graph data processing. She has published more than 100 papers in top venues such as TKDE, TODS, VLDBJ, SIGMOD, VLDB, and ICDE. She received the Discovery Early Career Researcher Award in 2011 and the prestigious Chris Wallace Award in 2019.</p> <p>Ying Zhang is an Australian ARC Future Fellow and Professor at the University of Technology, Sydney (UTS). He has been the head of the database group at the Centre for Artificial Intelligence (CAI) since 2014. His research focuses on efficient query processing and analytics on big data. He has published more than 80 papers in top venues. He had received seven ARC grants which are under the National Competitive Grants Programme (NCGP) including one ARC ADP fellowship, one ARC DECRA fellowship, one ARC future fellowship and four ARC discovery projects.</p>
<p>Tutorial # 3</p>	<p>Title: A tutorial on biomedical image segmentation using deep learning</p> <p>Speakers:</p>

<p>(April 13, 2022, Wednesday, 10.15-12.30, Hall 5)</p>	<ul style="list-style-type: none"> • Sonali Agarwal (IIIT Allahabad) • Krishna Pratap Singh (IIIT Allahabad) • Sanjay Kumar Sonbhadra (ITER Bhubaneswar) • Narinder Singh Punn (IIIT Allahabad) <p>Brief outline of the tutorial:</p> <p>Deep learning algorithms, in particular convolutional networks, have rapidly become a methodology of choice for analyzing medical images. Most medical applications require identifying and localizing the objects or regions (damaged tissues, cells or nuclei) found in medical imaging such as CAT scans, X-Rays, Ultrasound, etc. for diagnosis, monitoring and treatment. This delineation is generally performed by expert clinicians or radiologists which is a complex and time-consuming task. In recent studies, the implication of transfer learning and U-Net based approaches have illustrated state-of-the-art performance in different applications for the development of computer-aided diagnosis systems to localize the infected or damaged tissues or cells in the body using various modalities for early diagnosis and treatment of diseases such as brain tumor, lung cancer, Alzheimer, breast cancer, etc. With this motivation, this tutorial focuses on state-of-the-art deep learning approaches, a critical discussion of open challenges and directions for future research in the area of biomedical image segmentation.</p> <p>Biographies of Speakers:</p> <p>Sonali Agarwal is working as an Associate Professor in the Information Technology Department of Indian Institute of Information Technology (IIIT), Allahabad, India. She received her Ph. D. Degree at IIIT Allahabad and joined as faculty at IIIT Allahabad, where she has been teaching since October 2009. She holds Bachelor of Engineering (B.E.) degree in Electrical Engineering from Bhilai Institute of Technology, Bhilai, (C.G.) India and Masters of Engineering (M.E.) degree in Computer Science from Motilal Nehru National Institute of Technology (MNNIT), Allahabad, India Her main research interests are in the areas of Artificial Intelligence and Big Data. She is the head of Big Data Analytics Lab at IIIT Allahabad, India.</p> <p>Krishna Pratap Singh is working as an Associate professor in the Information Technology Department at IIIT Allahabad, India. He received his PhD and Master from IIT Roorkee. He has been working at IIIT Allahabad since 2009. His main research areas are Machine Learning, Transfer Learning and Optimization.</p> <p>Sanjay Kumar Sonbhadra is presently working as Assistant Professor in the Computer Science and Engineering Department of ITER, Shiksha ‘O’ Anusandhan, Bhubaneswar, Odisha, India. He is mainly working on One-class classification, Anomaly detection, Target class guided dimensionality reduction and training sample selection techniques and Big data analytics. During 2017-2021, he worked as a senior member of “Big Data Analytics Lab” at IIIT Allahabad, India. He has published many articles in the area of machine learning applications to address recent challenges of COVID-19. He has working experience of machine learning algorithms to address the challenging problem of target specific learning with limited target samples.</p> <p>Narinder Singh Punn is working as a Teaching Research Assistant (TRA) in the Information Technology Department of Indian Institute of Information Technology (IIIT), Allahabad, India. Narinder’s main research includes Medical Imaging segmentation, Deep learning and Artificial Intelligence techniques in healthcare. He is a senior member of “Big Data Analytics Lab” at IIIT Allahabad, India. His recent publications cover applications of deep learning in the detection and prevention of COVID-19, while also exploiting the potential of self-supervised learning in biomedical image segmentation.</p>
<p>Tutorial # 4 (April 13, 2022, Wednesday, 14.30-16.15, Hall 5)</p>	<p>Title: AI Meets NoSQL Database: Methods, Opportunities and Challenges.</p> <p>Speakers:</p> <ul style="list-style-type: none"> • Hongzhi Wang (Harbin Institute of Technology) • Zhixin Qi (Harbin Institute of Technology) • Yu Yan (Harbin Institute of Technology) <p>Brief outline of the tutorial:</p> <p>This tutorial is planned for 1.5 hours and consists of the following parts.</p> <p>(1) Background and Motivation (10’): We introduce the background of AI for database and motivate the need for applying AI techniques on NoSQL database with several scenarios.</p> <p>(2) Cost Estimation for Graph and Document Databases (20’): We discuss how AI techniques estimate query</p>

costs for graph and document databases.

(3) Physical Design for Key-Value and Graph Databases (20’): We introduce physical design methods for key-value and graph databases based on AI techniques.

(4) Index Recommendation for Key-Value and Document Databases (15’): We discuss existing AI-based index recommendation approaches for key-value and document databases.

(5) Opportunities and Challenges (20’): We present the research opportunities and challenges for NoSQL database management based on AI techniques.

(6) Summary (5’): We summarize this tutorial and give our critical thoughts to AI for NoSQL database.

Speakers Bio

Hongzhi Wang, Professor, PHD supervisor, the head of massive data computing center and the vice dean of the honors school of Harbin Institute of Technology, the secretary general of ACM SIGMOD China, outstanding CCF member, a standing committee member CCF databases and a member of CCF big data committee. Research Fields include big data management and analysis, database systems, knowledge engineering and data quality. He was “starring track” visiting professor at MSRA and postdoctoral fellow at University of California, Irvine. Prof. Wang has been PI for more than 10 national or international projects including NSFC key project, NSFC projects and National Technical support project, and co-PI for more than 10 national projects include 973 project, 863 project and NSFC key projects. He also serves as a member of ACM Data Science Task Force. He has won First natural science prize of Heilongjiang Province, MOE technological First award, Microsoft Fellowship, IBM PHD Fellowship and Chinese excellent database engineer. His publications include over 300 papers in the journals and conferences such as VLDB Journal, IEEE TKDE, VLDB, SIGMOD, ICDE and SIGIR, 6 books and 6 book chapters. His PHD thesis was elected to be outstanding PHD dissertation of CCF and Harbin Institute of Technology. He serves as the reviewer of more than 20 international journal including VLDB Journal, IEEE TKDE, and PC members of over 50 international conferences including SIGMOD, VLDB, KDD, ICML, NeurpIS, ICDE, etc. His papers were cited more than 3000 times. His personal website is <http://homepage.hit.edu.cn/wang>.

Zhixin Qi is currently a PhD student in School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China. She received her bachelor degree from Harbin Engineering University in 2016, and received her master degree from Harbin Institute of Technology in 2018. Her research interests include AI4DB, knowledge graph, and graph data management. She was awarded National Scholarship for PhD students in 2021, National Scholarship for master students in 2017, and National Scholarship for undergraduates in 2014. She has published more than 10 papers in international journals and conferences, including TKDE, KAIS, KBS, Neurocomputing, JCST, CIKM, and DASFAA.

Yu Yan is currently a PhD student in School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China. She received her master degree from Harbin Institute of Technology in 2021. She committed to the research of database tuning, multi-model database, and database auto-management. She got National Scholarship for master students in 2020. She has published many papers in international conferences and journals, such as Information Sciences, ApWeb and etc.

Tutorial # 5
(April 14,
2022,
Thursday,
10.15-12.30,
Hall 5)

Title: Time Series Anomaly Detection Toolkit for AI Applications

Speakers:

- Dhaval Patel (IBM Research)
- Dzung Phan (IBM Research)
- Markus Mueller (IBM Germany)

Brief outline of the tutorial:

The tutorial is organized in a sequence of three sections: Introduction, Theory and Hands-on-demo. In part one, we will briefly discuss foundations of time series dataset with the help of real-world examples. We will also present a broad taxonomy of time series dataset. We will also present general definition of anomalies in time series data and discuss three common variants of Anomaly/Outlier Detection problems. Next, we discuss basic machine learning primitives such as Estimator, Transformer, Data Stationarizer, etc that are useful for building anomaly pipeline. In machine learning field, these components become a backbone for building a complex model learning pipelines. We will formally introduce the key API such as “fit”, “predict”, “decision_function”, to the participant with

the help of 30+ different anomaly detection algorithms. Apart from provide the categorization of these algorithms, we will also discuss one algorithm namely Gaussian Graphical Model for interpretable anomaly detection. The access to the toolkit is made available via IBM API Hub Platform (<https://developer.ibm.com/apis/catalog/ai4industry--anomaly-detection-product/Introduction>). The example notebooks are accessible at IBM's public github (<https://github.com/IBM/anomaly-detection-code-pattern/>). The tutorial finally analyzes open issues and future directions in this vibrant and rapidly evolving research area.

Biographies of Speakers:

Dhaval Patel work at IBM Research since 2016. Dr. Dhaval Patel hold PhD in Computer Science from National University of Singapore and master's degree in Information Technology from Indian Institute of Technology – Kharagpur. Dr. Patel is an expert in Data Mining, Machine Learning, Time Series Data Analysis, etc. The significance of his research contributions has been demonstrated in 60+ published papers (10 journal papers and 50+ conference papers) in high impact, refereed, top-notch venues in Database, Data Mining, Big Data and Machine Learning, 1 issued US patent and 18 patent applications. He is recipients of 9 outstanding technical/research accomplishments awards from IBM for advancing AI technology to solve several real-world industrial problems. He is key contributor in many Flagship IBM Product including AutoAI-TS, Maximo Application Suites for Anomaly Detection at Scale, etc.

Dzung Phan is a Research Staff Member at IBM Research, New York, USA since 2010. He received a Ph.D. degree in applied mathematics from the University of Florida in 2010, a M.S. degree in computational engineering from National University of Singapore in 2004, and a B.S. degree in mathematics from Vietnam National University, Hanoi in 2001. His research interests include optimization theory and algorithms, machine learning, and operations research. In particular, he is currently working on anomaly and change detection, sparse learning, and data-driven decision making. He has published more than 40 technical papers in refereed conferences and journals including top machine learning/data mining conferences such as ICML, NeurIPS, IJCAI, and ICDM. He has also filed about 40 U.S. patents. He received the 2012 Pat Goldberg Best Paper Award and a 2020 INFORMS Wagner Prize semi-finalist.

Markus Muller studied Math, Operations Research and Computer Science and has over 2 decades of experience in IT in different roles, mainly as software architect. He has shifted towards Machine Learning in 2018, first as an architect for an NLU related offering, then as data scientist for an offering in the IIoT space. Markus worked at IBM Watson Center, Munich.

4.Workshops (April 11, 2022, Monday, 09.00- 18.00)

(All timings are as per Indian Standard Time (IST) (UTC + 05:30))

S,No	Workshop Name	Program
1.	1st Workshop on Pattern mining and Machine learning in Big complex Databases (PMBD 2021)	URL: https://philippe-fournier-viger.com/PMDB_2021/program.html
2.	The 6th International Workshop on Graph Data Management and	URL: https://gdma2022.github.io

	Analysis (GDMA 2022)	
3.	International Workshop on Blockchain Technologies (IWBT2022)	URL: https://sites.google.com/ahduni.edu.in/iwbt2022/home
4.	The 8th International Workshop on Big Data Management and Service (BDMS 2022)	URL: https://zheng-kai.com/cfp
5.	Managing Air Quality Through Data Science (MAQTDS 2022)	URL: https://ml2ct.ashoka.edu.in/en/events/dasfa-maqtlds/
6.	The 7th International Workshop on Big Data Quality Management (BDQM 2022)	URL: https://bdqm2022.github.io

5. Research sessions (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Research Session #	Session Name (Date, Day, Time, and Hall)	Paper ID	Paper Title
1	Research Session #1 (Queries), (April 12, 2022, Tuesday, 10.15-12.30, Hall 1)	20	Approximate Continuous Top-K Queries Over Memory Limitation-based Streaming Data
		22	Cross-Model Conjunctive Queries over Relation and Tree-structured Data
		178	Leveraging Search History for Improving Person-Job Fit
		503	Efficient In-Memory Evaluation of Reachability Graph Pattern Queries on Data Graphs
		523	Revisiting Approximate Query Processing and Bootstrap Error Estimation on GPU
		257	μ -join: Efficient Join with Versioned Dimension Tables
		532	Learning-based Optimization for Online Approximate Query Processing
2	Research Session #2 (Text and Image-I) (April 12, 2022, Tuesday, 10.15-12.30, Hall 2)	65	Emotion-aware Multimodal Pre-training for Image-grounded Emotional Response Generation
		416	KdTNet: Medical Image Report Generation via Knowledge-driven Transformer
		464	AdCSE: An Adversarial Method for Contrastive Learning of Sentence Embeddings
		502	HRG: A Hybrid Retrieval and Generation model in Multi-turn Dialogue Generation
		122	SimEmotion: A Simple Knowledgeable Prompt Tuning Method for Image Emotion Classification
		196	E-commerce Knowledge Extraction via Multi-modal Machine Reading Comprehension
		403	Concurrent Transformer for Spatial-Temporal Graph Modeling
		519	Modeling Uncertainty in Neural Relation Extraction
3		174	Similarity-Aware Collaborative Learning for Patient Outcome Prediction

	Research Session #3 (Applications of ML-I) (April 12, 2022, Tuesday, 10.15-12.30, Hall 3)	252	Heterogeneous Federated Learning via Grouped Sequential-to-Parallel Training
		335	Peripheral Instance Augmentation for End-to-End Anomaly Detection using Weighted Adversarial Learning
		342	HieNet: Bidirectional Hierarchy Framework for Automating ICD Coding
		2	Data Source Selection in Federated Learning: A Submodular Optimization Approach
		41	CLZT:A Contrastive Learning Based Framework for Zero-Shot Text Classification
		248	An Interpretable Time Series Classification Approach Based on Feature Clustering
		332	Supervised Multi-view Latent Space Learning by Jointly Preserving Similarities across Views and Samples
4	Research Session #4 (Graphs-I) (April 12, 2022, Tuesday, 14.45-16.30, Hall 1)	11	Cascade-Enhanced Graph Convolutional Network for Information Diffusion Prediction
		15	Diversify Search Results through Graph Attentive Document Interaction
		365	Learning Robust Representation through Graph Adversarial Contrastive Learning
		197	Contrastive Disentangled Graph Convolutional Network for Weakly-supervised Node Classification
		353	CSGNN: Improving Graph Neural Networks with Contrastive Semi-Supervised Learning
5	Research Session #5 (Text and Image-II) (April 12, 2022, Tuesday,14.45-16.30, Hall 2)	271	Tracking the Evolution: Discovering and Visualizing the Evolution of Literature
		323	Incorporating Commonsense Knowledge into Story Ending Generation via Heterogeneous Graph Networks
		400	Open-domain Dialogue Generation Grounded with Dynamic Multi-form Knowledge Fusion
		158	Knowing What I Don't Know: A Generation Assisted Rejection Framework in Knowledge Base Question Answering
		348	PromptMNER: Prompt-based Entity-related Visual Clue Extraction and Integration for Multimodal Named Entity Recognition
		423	Towards Personalized Review Generation with Gated Multi-source Fusion Network
		506	Definition-Augmented Jointly Training Framework for Intention Phrase Mining
6	Research Session #6 (Recommendation-I) (April 12, 2022, Tuesday, 14.45-16.30, Hall 3)	99	Inter- and Intra-Domain Relation-aware Heterogeneous Graph Convolutional Networks for Cross-Domain Recommendation
		294	Knowledge-Enhanced Multi-task Learning for Course Recommendation
		406	Joint Locality Preservation and Adaptive Combination for Graph Collaborative Filtering
		447	Hyperbolic Personalized Tag Recommendation
		14	Collaborative Filtering for Recommendation in Geometric Algebra
		397	SAER: Sentiment-opinion Alignment Explainable Recommendation
		437	Toward Auto-learning Hyperparameters for Deep Learning based Recommender Systems

7	Research Session #7 (Spatio-temporal Data) (April 13, 2022, Wednesday, 10.15-12.30, Hall 1)	200	JS-STDGN: A Spatial-Temporal Dynamic Graph Network using JS-Graph for Traffic Prediction
		311	When Multitask Learning Make a Difference: Spatio-temporal Joint Prediction for Cellular Trajectories
		7	Efficient Retrieval of Top-k Weighted Spatial Triangles
		47	DIOT: Detecting Implicit Obstacles from Trajectories
		267	Exploring Sub-skeleton Trajectories for Interpretable Recognition of Sign Language
		494	Significant Engagement Community Search on Temporal Networks
		510	Influence Computation for Indoor Spatial Objects
		516	NavFPNet: A Unified Framework for Network Localization in GPS-free Navigation Scenarios
8	Research Session #8 (Recommendation-II) (April 13, 2022, Wednesday, 10.15-12.30, Hall 2)	64	M ³ -IB: A Memory-augment Multi-modal Information Bottleneck Model for Next-item Recommendation
		75	Fully Utilizing Neighbors for Session-based Recommendation with Graph Neural Networks
		258	Enhancing Graph Convolution Network for Novel Recommendation
		301	PMAR: Multi-Aspect Recommendation Based on Psychological Gap
		133	Graph Neural Networks with Dynamic and Static Representations for Social Recommendation
		228	GISDCN: A Graph-based Interpolation Sequential Recommender with Deformable Convolutional Network
		359	Core Interests Focused Self-Attention for Sequential Recommendation
9	Research Session #9 (Applications of ML-II) (April 13, 2022, Wednesday, 10.15-12.30, Hall 3)	51	CaSS: A Channel-aware Self-supervised Representation Learning Framework for Multivariate Time Series Classification
		312	Transportation-mode Aware Travel Time Estimation via Meta-learning
		380	Efficient Consensus Motif Discovery of All Lengths in Multiple Time Series
		459	LiteWSC: a Lightweight Framework for Web-Scale Spectral Clustering
		40	MetisRL: A Reinforcement Learning Approach for Dynamic Routing in Data Center Networks
		128	Learning Evolving Concepts with Online Class Posterior Probability
		290	A Trace Ratio Maximization Method for Parameter Free Multiple Kernel Clustering
10	Research Session #10 (Algorithms-I) (April 13, 2022, Wednesday, 14.30-16.15, Hall 1)	181	Improving Information Cascade Modeling by Social Topology and Dual Role User Dependency
		277	Mining Negative Sequential Rules from Negative Sequential Patterns
		287	CrossIndex: Memory-Friendly And Session-Aware Index for Supporting Crossfilter in Interactive Data Exploration
		489	Hierarchical Bitmap Indexing for Range Queries on Multidimensional Arrays

		261	A Novel Null-invariant Temporal Measure to Discover Partial Periodic Patterns in Non-uniform Temporal Databases
		418	A Two-phase Approach for Recognizing Tables with Complex Structures
		467	A Dynamic Heterogeneous Graph Perception Network with Time-Based Mini-Batch for Information Diffusion Prediction
11	Research Session #11, (Systems) (April 13, 2022, Wednesday, 14.30-16.15, Hall 2)	310	HEM: A Hardware-aware Event Matching Algorithm for Content-based Pub/Sub Systems
		396	RotorRaft: Scalable Follower-Driven Raft on RDMA
		545	Efficient Matrix Computation for SGD-based Algorithms on Apache Spark
		204	Parallel Pivoted Subgraph Filtering with Partial Coding Trees on GPU
		317	Txchain: Scaling Sharded Decentralized Ledger via Chained Transaction Sequences
		471	Zebra: An Efficient, RDMA-Enabled Distributed Persistent Memory File System
12	Research Session #12 (Applications of ML-III) (April 13, 2022, Wednesday, 14.30-16. 15, Hall 3)	6	Hierarchical Attention Factorization Machine for CTR Prediction
		35	MCRF: Enhancing CTR Prediction Models via Multi-Channel Feature Refinement Framework
		185	Semi-Supervised Graph Learning with Few Labeled Nodes
		328	A Deep Reinforcement Learning Based Dynamic Pricing Algorithm in Ride-hailing
		87	InDISP: An Interpretable Model for Dynamic Illness Severity Prediction
		249	Generative Adversarial Imitation Learning to Search in Branch-and-Bound Algorithms
		373	Market-aware Dynamic Person-Job Fit with Hierarchical Reinforcement Learning
13	Research Session #13 (Security) (April 13, 2022, Wednesday, 16.30-18.00, Hall 1)	21	ADAPT: Adversarial Domain Adaptation with Purifier Training for Cross-domain Credit Risk Forecasting
		345	Poisoning Attacks on Fair Machine Learning
		284	Bi-Level Selection via Meta Gradient for Graph-based Fraud Detection
		344	Contrastive Learning for Insider Threat Detection
		355	Metadata Privacy Preservation for Blockchain-based Healthcare Systems
		384	Blockchain-Based Encrypted Image Storage and Search in Cloud Computing
14	Research Session #14 (Text and Image-III) (April 13, 2022, Wednesday, 16.30-18.00, Hall 2)	223	Semantic-based Data Augmentation for Math Word Problems
		455	Aligning Internal Regularity and External Influence of Multi-Granularity for Temporal Knowledge Graph Embedding
		518	FalCon: A Faithful Contrastive Framework for Response Generation in TableQA Systems
		80	Tipster: A Topic-Guided Language Model for Topic-Aware Text Segmentation
		155	Predicting Rumor Veracity on Social Media with Graph Structured Multi-task Learning
		188	KAAS: A Keyword-Aware Attention Abstractive Summarization Model for Scientific Articles

		198	PERM: Pre-training Question Embeddings via Relation Map for Improving Knowledge Tracing
15	Research Session #15 (Graphs-II) (April 13, 2022, Wednesday, 16.30-18.00, Hall 3)	98	On Glocal Explainability of Graph Neural Networks
		337	Temporal Network Embedding with Motif Structural Features
		392	What Affects the Performance of Models? Sensitivity Analysis of Knowledge Graph Embedding
		123	CollaborateCas: Popularity Prediction of Information Cascades based on Collaborative Graph Attention Networks
		356	IncreGNN: Incremental Graph Neural Network Learning by Considering Node and Parameter Importance
		427	Representation Learning in Heterogeneous Information Networks Based on Hyper Adjacency Matrix
16	Research Session #16 (Algorithms-II) (April 14, 2022, Thursday, 10.15-12.30, Hall 1)	194	Discovering Bursting Patterns over Streaming Graphs
		327	GHStore: A High Performance Global Hash based Key Value Store
		533	Membership Algorithm for Single-Occurrence Regular Expressions with Shuffle and Counting
		10	(p, n)-core: signed core decomposition in signed networks
		120	TROP: Task Ranking Optimization Problem on Crowdsourcing Service Platform
		240	HATree: A Hotness-Aware Tree Index with In-Node Hotspot Cache for NVM/DRAM-Based Hybrid Memory Architecture
		349	Utilizing Expert Knowledge and Contextual Information for Sample-limited Causal Graph Construction
		463	Towards Semantic correspondence between Statistical Reasoning, OLAP and Association Rule Mining: Semantics and Pragmatics
17	Research Session #17 (Recommendation-III) (April 14, 2022, Thursday, 10.15-12.30, Hall 2)	300	Learning Social Influence from Network Structure for Recommender Systems
		398	Meta-path Enhanced Lightweight Graph Neural Network for Social Recommendation
		402	Multi-view Multi-behavior Contrastive Learning in Recommendation
		452	Diffusion-based Graph Contrastive Learning for Recommendation with Implicit Feedback
		354	Deep Graph Mutual Learning for Cross-domain Recommendation
		453	GELibRec: Third-Party Libraries Recommendation Using Graph Neural Network
18	Research Session #18 (Applications of ML-IV) (April 14, 2022, Thursday, 10.15-12.30, Hall 3)	135	Temporal Knowledge Graph Entity Alignment via Representation Learning
		208	Human Mobility Identification by Deep Behavior Relevant Location Representation
		469	Dual Confidence Learning Network for Open-World Time Series Classification
		530	Port Container Throughput Prediction Based on Variational AutoEncoder
		150	Robust Dynamic Pricing in Online Markets with Reinforcement Learning
		246	Multi-Memory enhanced Separation Network for Indoor Temperature Prediction

		422	TEALED: A Multi-Step Workload Forecasting Approach Using Time-Sensitive EMD and Auto LSTM Encoder-Decoder
19	Research Session #19 (Knowledge Bases) (April 14, 2022, Thursday, 14.30-16.15, Hall 1)	44	Triple-as-Node Knowledge Graph and Its Embeddings
		170	LeKAN: Extracting Long-tail Relations via Layer-Enhanced Knowledge-Aggregation Networks
		264	TRHyTE: Temporal Knowledge Graph Embedding based on Temporal-Relational Hyperplanes
		168	ExKGR: Explainable Multi-hop Reasoning for Evolving Knowledge Graph
		243	Improving Core Path Reasoning for the Weakly Supervised Knowledge Base Question Answering
		360	Counterfactual-guided and Curiosity-driven Multi-Hop Reasoning over Knowledge Graph
		388	Visualizable or Non-visualizable? Exploring the Visualizability of Concepts in Multi-modal Knowledge Graph
20	Research Session #20 (Text and Image-IV) (April 14, 2022, Thursday, 14.30-16.15, Hall 2)	90	Information Networks based Multi-semantic Data Embedding for Entity Resolution
		269	Empowering Transformer with Hybrid Matching Knowledge for Entity Matching
		445	Fake Restaurant Review Detection Using Deep Neural Networks with Hybrid Feature Fusion Method
		179	Medical image fusion based on pixel-level nonlocal self-similarity prior and optimization
		187	Knowledge-enhanced Interactive Matching Network for Multi-turn Response Selection in Medical Dialogue Systems
		268	A Three-Stage Curriculum Learning Framework with Hierarchical Label Smoothing for Fine-Grained Entity Typing
		395	TaskSum: Task-Driven Extractive Text Summarization for Long News Documents Based on Reinforcement Learning
21	Research Session #21 (Recommendation-IV) (April 14, Thursday, 2022, 14.30-16.15, Hall 3)	12	MDKE: Multi-level Disentangled Knowledge-Based Embedding for Recommender Systems
		401	Intention Adaptive Graph Neural Network for Category-aware Session-based Recommendation
		430	Gated Hypergraph Neural Network for Scene-aware Recommendation
		13	Multi-Behavior Recommendation with Two-Level Graph Attentional Networks
		157	Toward Paper Recommendation by Jointly Exploiting Diversity and Dynamics in Heterogeneous Information Networks
		166	Enhancing Session-based Recommendation with Global Context Information and Knowledge Graph

6. Industry Sessions (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Industry Session Id	Session Name (Date, Time, and Hall)	Paper ID	Paper Title
1	Industry Session # 1 Advancing Recommendation Systems (April 12, 2022, Tuesday, 10.15-12.30, Hall 4)	553	A Joint Framework for Explainable Recommendation with Knowledge Reasoning and Graph Representation Authors: Zhang, Luhao; Fang, Ruiyu; Yang, Tianchi; Hu, Maodi; Shi, Chuan; Li, Tao; Wang, Dong
		568	XDM: Improving Sequential Deep Matching with Unclicked User Behaviors for Recommender System Authors: Lv, Fuyu; Li, Mengxue; Guo, Tonglei; Yu, Changlong; Sun, Fei; Jin, Taiwei; Ng, Wilfred
		584	Mitigating Popularity Bias in Recommendation via Counterfactual Inference (Industry Track) Ming He Authors: He, Ming; Li, Changshu; Hu, Xinlei; Chen, Xin; Wang, Jiwen
		594	Efficient Dual-process Cognitive Recommender Balancing Accuracy and Diversity Zhongyu Wei Authors: Gao, Yixu; Shao, Kun; Duan, Zhijian; Wei, Zhongyu*; Li, Dong; Wang, Bin; Zhao, Mengchen; Hao, Jianye
		629	Learning and Fusing Multiple User Interest Representations for Sequential Recommendation Authors: He, Ming; Han, Tianshuo; Ding, Tianyu
2	Industry Session # 2 Data Management and Search (April 13, 2022, Wednesday, 10.15-12.30, Hall 4)		Invited talk Title: DocLens: Digiizing healthcare documents By Gaurav Aggarwal, Sujoy Paul, Narayan Hegde Google
		627	Query-Document Topic Mismatch Detection Manish Gupta Authors: Chelaramani, Sahil; Chatterjee, Ankush; Damani, Sonam; Narahari, Kedhar Nath; Joshi, Meghana; Gupta, Manish; Agrawal, Puneet
		561	Beyond QA: 'Heuristic QA' Strategies in JIMI Authors: Song, Shuangyong; Zou, Bo; Lin, Jianghua; Yu, Xiaoguang; He, Xiaodong
		571	SQLG+: Efficient k -hop Query Processing on RDBMS Li Zeng Authors: Zeng, Li; Zhou, Jinhua; Qin, Shijun; Cai, Haoran; Zhao, Rongqian; Chen, Xin
3	Industry Session # 3 Industrial Machine Learning Applications (April 14, 2022, Thursday, 10.15-12.30, Hall 4)	570	Modeling Long-Range Travelling Times with Big Railway Data Authors: Sun, Wenya; Grubenman, Tobias; Cheng, Reynold; Kao, Ben; Ching, Wai-Ki
		579	Multi-scale Time Based Stock Appreciation Ranking Prediction via Price Co-movement Discrimination Authors: Xu, Ruyao; Cheng, Dawei; Chen, Cen; Luo, Siqiang; Luo, Yifeng; Qian, Weining
		591	RShield: A Refined Shield for Complex Multi-Step Attack Detection based on Temporal Graph Network

		Authors: Yang, Weiyong; Gao, Peng; Huang, Hao; Wei, Xingshen; Liu, Wei; Zhu, Shishun; Luo, Wang
	592	Inter-and-Intra Domain Attention Relational Inference for Cabinet Temperature Prediction in Data Center Authors: Shen, Fang; Li, Zhan; Pan, Bing; Zhang, Ziwei; Wang, Xin; Wang, Jialong; Zhao, Wendy; Zhu, Wenwu

7. Demo and PhD Consortium Session (April 12, 2022, 13.30-14.45, Tuesday, Hall 1), (All timings are as per Indian Standard Time (IST) (UTC + 05:30))

Session Name	Paper ID	Paper Title
Demo and PhD Consortium	598	An Interactive Data Imputation System Authors: Yangyang Wu (Zhejiang University); Xiaoye Miao (Zhejiang University); Yuchen Peng (Zhejiang University); Lu Chen (Zhejiang University); Yunjun Gao (Zhejiang University); Jianwei Yin (Zhejiang University).
	599	FoodChain: A Food Delivery Platform Based on Blockchain for Keeping Data Privacy Authors: Rodrigo B Folha (Federal University of Pernambuco); Valéria Times (Federal University of Pernambuco); Arthur Carvalho (Miami University); André Araújo (Federal University of Alagoas); Henrique Couto (Federal University of Alagoas); Flaviano Viana (Federal University of Pernambuco).
	601	A scalable lightweight RDF knowledge retrieval system Authors: Yuming Lin (Guilin University of Electronic Technology); Chuangxin Fang (Guilin University of Electronic Technology).
	603	CO-AutoML: An Optimizable Automated Machine Learning System Authors: Chunnan Wang (HIT); Hongzhi Wang (Harbin Institute of Technology); Xu Bo (Harbin Institute of Technology); Xintong Song (Harbin Institute of Technology); Xiangyu Shi (Harbin Institute of Technology); Yuhao Bao (Harbin Institute of Technology).
	604	OIIKM: A System for Discovering Implied Knowledge from Spatial Datasets Using Ontology Authors: Long Wang (Guilin University Of Electronic Technology)*; Xuguang Bao (Guilin University of Electronic Technology); Liang Chang (Guilin University of Electronic Technology); Tianlong Gu (Guilin University Of Electronic Technology).
	608	IDMBS: An Interactive System to Find Interesting Co-location Patterns Using SVM Authors: Yuxiang Zhang (Guilin University of Electronic Technology)*; Xuguang Bao (Guilin University of Electronic Technology); Liang Chang (Guilin University of Electronic Technology); Tianlong Gu (Guilin University Of Electronic Technology).
	609	A Secure Trajectory Similarity Search System Authors: Yiping Teng (Shenyang Aerospace University); Fanyou Zhao (Shenyang Aerospace University); Jiayv Liu (Shenyang Aerospace University); Mengfan Zhang (Shenyang Aerospace University); Jihang Duan (Shenyang Aerospace University); Zhan Shi (Shenyang Aerospace University).

	612	Data-based Insights for the Masses: Scaling Natural Language Querying to Middleware Data Authors: Kausik Lakkaraju (University of South Carolina)*; Vinamra Palaiya (Tantiv4); Sai Teja Paladi (University of South Carolina); Chinmayi Appajigowda (Tantiv4); Biplav Srivastava (AI Institute, University of South Carolina); Lokesh Johri (tantiv4).
	615	Identifying relevant sentences for travel blogs from Wikipedia articles Authors: Arnav Kapoor (International Institute of Information Technology, Hyderabad)*; Manish Gupta (Microsoft).
PhD Consortium presentations	630	Neuro-Symbolic XAI for Computational Drug Repurposing Author: Martin Drancé Martin Drancé (University of Bordeaux)
	639	Leveraging Non-negative Matrix Factorization for Document Summarization Author: Alka Khurana (Department of Computer Science, University of Delhi)

*****PROGRAM END HERE*****